

Rec'd PCT/PTO 01 APR 2005

10/528674



PCT

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,674

TIME: 10:54:06

Input Set : A:\0185659SeqList.txt

Output Set: N:\CRF4\04012005\J528674.raw

3 <110> APPLICANT: DSM IP ASSETS B.V.

5 <120> TITLE OF INVENTION: Transcriptional activator gene for genes involved in cobalamin biosynthesis

6 biosynthesis

8 <130> FILE REFERENCE: NDR5233

C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/528,674

C--> 11 <141> CURRENT FILING DATE: 2005-03-23

13 <150> PRIOR APPLICATION NUMBER: EP 02021603.2

14 <151> PRIOR FILING DATE: 2002-09-27

16 <160> NUMBER OF SEQ ID NOS: 8

18 <170> SOFTWARE: PatentIn version 3.2

20 <210> SEQ ID NO: 1

21 <211> LENGTH: 1439

22 <212> TYPE: DNA

23 <213> ORGANISM: Pseudomonas denitrificans

26 <220> FEATURE:

27 <221> NAME/KEY: CDS

28 <222> LOCATION: (303)..(1304)

30 <400> SEQUENCE: 1

31 cttttaccga caccgcacc agccgtgtga agcgtcgcaa tcggcaaaca caaccctgga 60

33 gacgcgtttg cgaatgttgc gtatctctgg cagttcttgt ccgatagcgt caacagcgct 120

35 gtaaaaaactg tcgcttttct taaaaagcc tgaaagcggg gccgagccgc ctttcaccgc 180

37 gcgatgtcgc aatgcgaaat ctctttgcgg tcgctgctat ccatgcgaat gtcgcaaaca 240

39 gacaagccga agagtatctt ccaagcgacg aatataccgc atcgccaagc atgggaagcc 300

41 tc atg aac aag cca ctg acc aaa aag cgt tct ctc gtc ttc ttc ctg 347

42 Met Asn Lys Pro Leu Thr Lys Lys Arg Ser Leu Val Phe Phe Leu

43 1 5 10 15

45 gtg ccg aac ttt tcc atg ctg ccc ttt tcg gcg gcg atc gaa acg ctc 395

46 Val Pro Asn Phe Ser Met Leu Pro Phe Ser Ala Ala Ile Glu Thr Leu

47 20 25 30

49 cgc atc gcc aac cgc atg ctc ggc tac gag gcc tat tcc tgg cgc ctc 443

50 Arg Ile Ala Asn Arg Met Leu Gly Tyr Glu Ala Tyr Ser Trp Arg Leu

51 35 40 45

53 gca tcg tcc gac ggc gaa aag gtc ctg tcg tcg agc ggt atc gcg ctc 491

54 Ala Ser Ser Asp Gly Glu Lys Val Leu Ser Ser Ser Gly Ile Ala Leu

55 50 55 60

57 gag gtc aac tcg tcg ctt gca gac gag cgc aag ttt ctc ggc ggc gaa 539

58 Glu Val Asn Ser Ser Leu Ala Asp Glu Arg Lys Phe Leu Gly Gly Glu

59 65 70 75

61 aac cgc ccc tcg atg gtg ctg gtc tgt tcc ggc atc tat gtc gag gac 587

62 Asn Arg Pro Ser Met Val Leu Val Cys Ser Gly Ile Tyr Val Glu Asp

63 80 85 90 95

65 ttc aac aac aag tcg gtc aat gcc tgg ctg cgc gag gtc tac aat cgc 635

66 Phe Asn Asn Lys Ser Val Asn Ala Trp Leu Arg Glu Val Tyr Asn Arg

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,674

TIME: 10:54:06

Input Set : A:\0185659SeqList.txt

Output Set: N:\CRF4\04012005\J528674.raw

```

67          100          105          110
69 ggc gtc gcc gtc ggc agc ctc tgt acc ggc gcc cat gtg ctg gcg tcg      683
70 Gly Val Ala Val Gly Ser Leu Cys Thr Gly Ala His Val Leu Ala Ser
71          115          120          125
73 gcc ggt ctt ctg acc ggc aag cgc tgc gcc atc cac tgg gaa aac ctg      731
74 Ala Gly Leu Leu Thr Gly Lys Arg Cys Ala Ile His Trp Glu Asn Leu
75          130          135          140
77 ccg ggc ttt tcc gaa agc ttc ccg cag gtc gac gtc tat gcc gac ctc      779
78 Pro Gly Phe Ser Glu Ser Phe Pro Gln Val Asp Val Tyr Ala Asp Leu
79          145          150          155
81 tac gaa atc gac agc aac atc tac acc tgc gcc ggc ggc acc gcc tcg      827
82 Tyr Glu Ile Asp Ser Asn Ile Tyr Thr Cys Ala Gly Gly Thr Ala Ser
83 160          165          170          175
85 ctc gac atg atg ctg aac ctg atc gac cag gat ttc ggc gag agc ctc      875
86 Leu Asp Met Met Leu Asn Leu Ile Asp Gln Asp Phe Gly Glu Ser Leu
87          180          185          190
89 gtc aac cgc gtc tgc gaa cag gcg ctg acc gat cgc gtg cgc ggg ccc      923
90 Val Asn Arg Val Cys Glu Gln Ala Leu Thr Asp Arg Val Arg Gly Pro
91          195          200          205
93 cat gac cgc cag cgc ctg ccg ctg cgc gcc cgt ctc ggc gtg cag aac      971
94 His Asp Arg Gln Arg Leu Pro Leu Arg Ala Arg Leu Gly Val Gln Asn
95          210          215          220
97 gcc aag gtg ctg tcc atc atc gaa ctg atg gag gca aac ctc gcc gag      1019
98 Ala Lys Val Leu Ser Ile Ile Glu Leu Met Glu Ala Asn Leu Ala Glu
99          225          230          235
101 ccg ctt tcg ctg ctc gaa atc gcc gag ggc gcc gat ctc tcc cgc cgc      1067
102 Pro Leu Ser Leu Leu Glu Ile Ala Glu Gly Ala Asp Leu Ser Arg Arg
103 240          245          250          255
105 cag atc gag cgc ctc ttc cgc cag gaa atg ggc cgc tcg cct gca cgc      1115
106 Gln Ile Glu Arg Leu Phe Arg Gln Glu Met Gly Arg Ser Pro Ala Arg
107          260          265          270
109 tac tat ctc gaa atc cgc ctc gat cgc gca agg cac ctc ttg atc cag      1163
110 Tyr Tyr Leu Glu Ile Arg Leu Asp Arg Ala Arg His Leu Leu Ile Gln
111          275          280          285
113 tcg tcg atg ccg gtg gtc gaa gtg gcc gta gcc tgc ggc ttc gtc tcc      1211
114 Ser Ser Met Pro Val Val Glu Val Ala Val Ala Cys Gly Phe Val Ser
115          290          295          300
117 gcc tcg cac ttc tcc aag tgt tat cgc gaa ctc tac aac cgc tcg ccg      1259
118 Ala Ser His Phe Ser Lys Cys Tyr Arg Glu Leu Tyr Asn Arg Ser Pro
119          305          310          315
121 cag cag gag cgc gcc gac cgc aag ctg acg ctg cag atg gcg cga      1304
122 Gln Gln Glu Arg Ala Asp Arg Lys Leu Thr Leu Gln Met Ala Arg
123 320          325          330
125 taagcggcag atcagatgga tcagacaagg cggagcttct ccgccctttt tcgttggtga      1364
127 cactttccgc ttgtgccgtc ctggcgcttg cacgcgccgc tgttttggat tgaatggtcg      1424
129 gcgtttcagg agagc      1439
132 <210> SEQ ID NO: 2
133 <211> LENGTH: 334
134 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,674

TIME: 10:54:06

Input Set : A:\0185659SeqList.txt

Output Set: N:\CRF4\04012005\J528674.raw

```

135 <213> ORGANISM: Pseudomonas denitrificans
137 <400> SEQUENCE: 2
139 Met Asn Lys Pro Leu Thr Lys Lys Arg Ser Leu Val Phe Phe Leu Val
140 1 5 10 15
143 Pro Asn Phe Ser Met Leu Pro Phe Ser Ala Ala Ile Glu Thr Leu Arg
144 20 25 30
147 Ile Ala Asn Arg Met Leu Gly Tyr Glu Ala Tyr Ser Trp Arg Leu Ala
148 35 40 45
151 Ser Ser Asp Gly Glu Lys Val Leu Ser Ser Ser Gly Ile Ala Leu Glu
152 50 55 60
155 Val Asn Ser Ser Leu Ala Asp Glu Arg Lys Phe Leu Gly Gly Glu Asn
156 65 70 75 80
159 Arg Pro Ser Met Val Leu Val Cys Ser Gly Ile Tyr Val Glu Asp Phe
160 85 90 95
163 Asn Asn Lys Ser Val Asn Ala Trp Leu Arg Glu Val Tyr Asn Arg Gly
164 100 105 110
167 Val Ala Val Gly Ser Leu Cys Thr Gly Ala His Val Leu Ala Ser Ala
168 115 120 125
171 Gly Leu Leu Thr Gly Lys Arg Cys Ala Ile His Trp Glu Asn Leu Pro
172 130 135 140
175 Gly Phe Ser Glu Ser Phe Pro Gln Val Asp Val Tyr Ala Asp Leu Tyr
176 145 150 155 160
179 Glu Ile Asp Ser Asn Ile Tyr Thr Cys Ala Gly Gly Thr Ala Ser Leu
180 165 170 175
183 Asp Met Met Leu Asn Leu Ile Asp Gln Asp Phe Gly Glu Ser Leu Val
184 180 185 190
187 Asn Arg Val Cys Glu Gln Ala Leu Thr Asp Arg Val Arg Gly Pro His
188 195 200 205
191 Asp Arg Gln Arg Leu Pro Leu Arg Ala Arg Leu Gly Val Gln Asn Ala
192 210 215 220
195 Lys Val Leu Ser Ile Ile Glu Leu Met Glu Ala Asn Leu Ala Glu Pro
196 225 230 235 240
199 Leu Ser Leu Leu Glu Ile Ala Glu Gly Ala Asp Leu Ser Arg Arg Gln
200 245 250 255
203 Ile Glu Arg Leu Phe Arg Gln Glu Met Gly Arg Ser Pro Ala Arg Tyr
204 260 265 270
207 Tyr Leu Glu Ile Arg Leu Asp Arg Ala Arg His Leu Leu Ile Gln Ser
208 275 280 285
211 Ser Met Pro Val Val Glu Val Ala Val Ala Cys Gly Phe Val Ser Ala
212 290 295 300
215 Ser His Phe Ser Lys Cys Tyr Arg Glu Leu Tyr Asn Arg Ser Pro Gln
216 305 310 315 320
219 Gln Glu Arg Ala Asp Arg Lys Leu Thr Leu Gln Met Ala Arg
220 325 330
223 <210> SEQ ID NO: 3
224 <211> LENGTH: 20
225 <212> TYPE: DNA
226 <213> ORGANISM: Artificial Sequence
228 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,674

TIME: 10:54:06

Input Set : A:\0185659SeqList.txt

Output Set: N:\CRF4\04012005\J528674.raw

```

229 <223> OTHER INFORMATION: Primer N-GUS
231 <400> SEQUENCE: 3
232 ccaggaatag gcctcgtagc 20
235 <210> SEQ ID NO: 4
236 <211> LENGTH: 31
237 <212> TYPE: DNA
238 <213> ORGANISM: Artificial Sequence
240 <220> FEATURE:
241 <223> OTHER INFORMATION: Primer C-GUS
243 <400> SEQUENCE: 4
244 cgcagagctc gagccaccga ggctgtagcc g 31
247 <210> SEQ ID NO: 5
248 <211> LENGTH: 32
249 <212> TYPE: DNA
250 <213> ORGANISM: Artificial Sequence
252 <220> FEATURE:
253 <223> OTHER INFORMATION: oligonucleotide 1 for cassette
255 <400> SEQUENCE: 5
256 ggatcctgac taactagcat ctggaggaca ca 32
259 <210> SEQ ID NO: 6
260 <211> LENGTH: 38
261 <212> TYPE: DNA
262 <213> ORGANISM: Artificial Sequence
264 <220> FEATURE:
265 <223> OTHER INFORMATION: oligonucleotide 2 for cassette
267 <400> SEQUENCE: 6
268 tatgtgtcct ccagatgcta gttagtcagg atcctgca 38
271 <210> SEQ ID NO: 7
272 <211> LENGTH: 25
273 <212> TYPE: DNA
274 <213> ORGANISM: Artificial Sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: Primer COBRNde
279 <400> SEQUENCE: 7
280 ggcatatgaa caagccactg accaa 25
283 <210> SEQ ID NO: 8
284 <211> LENGTH: 34
285 <212> TYPE: DNA
286 <213> ORGANISM: Artificial Sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Primer COBR135R
291 <400> SEQUENCE: 8
292 gtcacccggg catatgttac gtcctgtaga aacc 34

```

VERIFICATION SUMMARY

PATENT APPLICATION: **US/10/528,674**

DATE: 04/01/2005

TIME: 10:54:07

Input Set : **A:\0185659SeqList.txt**

Output Set: **N:\CRF4\04012005\J528674.raw**

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date